On the Trace and the Sum of Elements of a Matrix

Jorma Kaarlo Merikoski
Department of Mathematical Sciences
University of Tampere
P.O. Box 607
SF-33101 Tampere 10, Finland

Submitted by Emeric Deutsch

ABSTRACT

It is demonstrated that in many situations the sum of elements and the trace of a matrix behave similarly.

---

1. INTRODUCTION

We use the following notation:

$n \in \mathbb{Z}_+$, fixed;
$\mathscr{M}^n$ = the space of real $n \times n$ matrices;
$\mathscr{M}^n_+$ = the cone of real and nonnegative $n \times n$ matrices;
$\mathscr{S}^n$ = the space of real and symmetric $n \times n$ matrices;
$\mathscr{F}^n$ = the cone of real, symmetric, and nonnegative definite $n \times n$ matrices;
$A = (a_{ik}) \in \mathscr{M}^n$;
$\operatorname{su} A = \sum_i \sum_k a_{ik}$;
$\operatorname{tr} A = \sum_i a_{ii}$;
$\rho(A)$ = the spectral radius of A;
$\{\lambda_1, \ldots, \lambda_n\}$ = the spectrum of A; $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ if the eigenvalues are real;
$\{U_1, \ldots, U_n\}$ = an orthonormal basis of corresponding eigenvectors (if one exists);
$E = (1, \ldots, 1)^T \in \mathbb{R}^n$;
$E = c_1 U_1 + \cdots + c_n U_n$.

Let $m \in \mathbb{Z}_+$. The formal analogy of

$$\operatorname{su} A^m = E^T A^m E = c_1^2 \lambda_1^m + \cdots + c_n^2 \lambda_n^m \tag{1}$$

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

with

$$\operatorname{tr} A^m = \lambda_1^m + \cdots + \lambda_n^m \tag{2}$$

motivates us to study to what extent the properties of su and tr are similar. We are especially interested in whether tr has properties analogous to those of su presented in [1], [6], [7], [8], [9]. Many of our results are well known, but it may be of interest to systematize them.

2. ELEMENTARY PROPERTIES

The following propositions are well known and/or easy to prove.

PROPOSITION 1 (Linearity).

(S_1) $\operatorname{su}(A+B) = \operatorname{su} A + \operatorname{su} B$, $\operatorname{su}(cA) = c \operatorname{su} A$, $\forall A, B \in \mathcal{M}^n$, $c \in \mathbb{R}$;
(T_1) $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$, $\operatorname{tr}(cA) = c \operatorname{tr} A$, $\forall A, B \in \mathcal{M}^n$, $c \in \mathbb{R}$.

PROPOSITION 2 (Definiteness).

(S_2) $\operatorname{su} A \geqslant 0 \wedge (\operatorname{su} A = 0 \Leftrightarrow A = 0)$, $\forall A \in \mathcal{M}_+^n$;
(T_2) $\operatorname{tr} A \geqslant 0 \wedge (\operatorname{tr} A = 0 \Leftrightarrow A = 0)$, $\forall A \in \mathcal{F}^n$.

Thus su *is a norm in* $\mathcal{M}_+^n$, *and* tr *in* $\mathcal{F}^n$.

PROPOSITION 3 (Inner product properties).

(S_3) $\langle A, B \rangle = \operatorname{su} A^T B$ *is a semi-inner-product in* $\mathcal{M}^n$ *and an inner product in* $\mathcal{M}_+^n$;
(T_3) $\langle A, B \rangle = \operatorname{tr} A^T B$ *is an inner product in* $\mathcal{M}^n$.

Thus $(\operatorname{su} A^T A)^{1/2} = [\Sigma_i(\Sigma_k a_{ik})^2]^{1/2}$ *is a seminorm in* $\mathcal{M}^n$ *and a norm in* $\mathcal{M}_+^n$, *and* $(\operatorname{tr} A^T A)^{1/2} = (\Sigma_i \Sigma_k a_{ik}^2)^{1/2}$ *is a norm in* $\mathcal{M}^n$, *the well-known Frobenius norm.*

PROPOSITION 4 (Cauchy-Schwarz inequality).

(S_4) $|\operatorname{su} A^T B| \leqslant (\operatorname{su} A^T A)^{1/2} (\operatorname{su} B^T B)^{1/2}$, $\forall A, B \in \mathcal{M}^n$;
(T_4) $|\operatorname{tr} A^T B| \leqslant (\operatorname{tr} A^T A)^{1/2} (\operatorname{tr} B^T B)^{1/2}$, $\forall A, B \in \mathcal{M}^n$.

PROPOSITION 5 (Submultiplicativity).

(S_5) $\operatorname{su} AB \leqslant \operatorname{su} A \operatorname{su} B$, $\forall A, B \in \mathscr{M}_+^n$;
(T_5) $|\operatorname{tr} AB| \leqslant \operatorname{tr} A \operatorname{tr} B$, $\forall A, B \in \mathscr{F}^n$.

3. SUM AND TRACE OF POWERS

Let μ be a norm in $\mathscr{M}^n$. The following is well known (see e.g. [10], [11]): If μ is submultiplicative, then

(N_1) $\rho(A) \leqslant \mu(A)$, $\forall A \in \mathscr{M}^n$;

and for any μ,

(N_2) $\lim_{m \to \infty} \mu(A^m)^{1/m} = \rho(A)$, $\forall A \in \mathscr{M}^n$.

From (N_2) it follows easily that

(N_3) $\lim_{m \to \infty} \dfrac{\mu(A^{m+1})}{\mu(A^m)} = \rho(A)$

if the limit exists.

We study whether su and tr have properties related to (N_1), (N_2), (N_3).

PROPOSITION 6 (Spectral dominance).

(S_6) $\rho(A) \leqslant \operatorname{su} A$, $\forall A \in \mathscr{M}_+^n$;
(T_6) $\rho(A) \leqslant \operatorname{tr} A$, $\forall A \in \mathscr{F}^n$.

PROPOSITION 7.

(S_7) $\lim_{m \to \infty} (\operatorname{su} A^m)^{1/m} = \left\{ \begin{matrix} \rho(A), & \forall A \in \mathscr{M}_+^n \\ \lambda_p, & \forall A \in \mathscr{F}^n \end{matrix} \right\}$ *with*

$$c_1 = \cdots = c_{p-1} = 0, \qquad c_p \neq 0; \tag{3}$$

(T_7) $\lim_{m \to \infty} (\operatorname{tr} A^m)^{1/m} = \rho(A)$, $\forall A \in \mathscr{M}^n$, *with*

$$\lambda_1 > |\lambda_n|. \tag{4}$$

PROPOSITION 8.

(S_8) $\lim_{m\to\infty} \dfrac{\operatorname{su} A^{m+1}}{\operatorname{su} A^m} = \lambda_p$, $\forall A \in \mathscr{F}^n \setminus \{0\}$ *satisfying* (3) (*cf.* [8, *Theorem* 3]);

(T_8) $\lim_{m\to\infty} \dfrac{\operatorname{tr} A^{m+1}}{\operatorname{tr} A^m} = \rho(A)$, $\forall A \in \mathscr{M}^n \setminus \{0\}$ *satisfying* (4).

PROPOSITION 9. *The sequence* (x_m) *is*

(S_9) (a) *decreasing if* $x_m = (\operatorname{su} A^m)^{1/m}$, $\forall A \in \mathscr{M}_+^n \cap \mathscr{F}^n$;

(b) *increasing if* $x_m = \left(\dfrac{\operatorname{su} A^m}{n}\right)^{1/m}$, $\forall A \in \mathscr{F}^n$ (*see* [6, *p.* 518]);

(c) *increasing if* $x_m = \dfrac{\operatorname{su} A^{m+1}}{\operatorname{su} A^m}$, $\forall A \in \mathscr{F}^n \setminus \{0\}$ (*cf.* [8, *Theorem* 3]);

(T_9) (a) *decreasing if* $x_m = (\operatorname{tr} A^m)^{1/m}$, $\forall A \in \mathscr{F}^n$;

(b) *increasing if* $x_m = \left(\dfrac{\operatorname{tr} A^m}{n}\right)^{1/m}$, $\forall A \in \mathscr{F}^n$;

(c) *increasing if* $x_m = \dfrac{\operatorname{tr} A^{m+1}}{\operatorname{tr} A^m}$, $\forall A \in \mathscr{F}^n \setminus \{0\}$.

Proof of Proposition 9. By (1) and (2), (T_9)(a) follows from Jensen's inequality [2, p. 28], and (S_9)(b) and (T_9)(b) from Schlömlich's inequality [2, p. 26]. (S_9)(c) and (T_9)(c) can be shown by straightforward differentiation.

(S_9)(a) remains to be proved; we use the same idea as in the proof of Jensen's inequality. Omitting the trivial case $A = 0$, we can restrict $\lambda_1 = 1$. By the Perron-Frobenius theory, there exists $U_1 \geqslant 0$ (elementwise); hence $c_1 = E^T U_1 \geqslant E^T(1,0,\dots,0)^T = 1$. Now

$$x_m = \left(c_1^2 + c_2^2\lambda_2^m + \cdots + c_n^2\lambda_n^m\right)^{1/m} = K_m^{1/m}$$

with $1 \geqslant \lambda_2,\dots,\lambda_n \geqslant 0$, $K_m \geqslant 1$. As m increases, $\lambda_2^m,\dots,\lambda_n^m$ decrease; hence K_m decreases, being $\geqslant 1$. Thus also x_m decreases. ∎

Nonnegativity or nonnegative definiteness of $A \in \mathscr{S}^n$ alone does not guarantee (S_9)(a). For example, let

$$A = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}, \qquad \operatorname{su} A = 1.$$

Then $A \in \mathscr{F}^2$, but

$$A^2 = \begin{pmatrix} 20 & -10 \\ -10 & 5 \end{pmatrix}, \qquad (\operatorname{su} A^2)^{1/2} = 5^{1/2}.$$

For the second example, consider

$$A = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \in \mathscr{M}_+^n. \tag{5}$$

Then

$$A^2 = \begin{pmatrix} 1 & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & n-1 \end{pmatrix}, \qquad A^3 = (n-1)A, \qquad A^4 = (n-1)A^2.$$

Hence

$$(\mathrm{su}\, A^3)^{1/3} = 2^{1/3}(n-1)^{2/3}, \qquad (\mathrm{su}\, A^4)^{1/4} = n^{1/4}(n-1)^{1/2},$$

the former being less than the latter if $n \geqslant 14$. The question of finding a counterexample $\in \mathscr{S}^n \cap \mathscr{M}_+^n$ for $n < 14$ remains open.

It is easy to see (cf. Proposition 10) that

$$(\mathrm{su}\, A^3)^{1/3} \leqslant (\mathrm{su}\, A^2)^{1/2} \leqslant \mathrm{su}\, A, \qquad \forall A \in \mathscr{M}_+^n.$$

Neither does $A \in \mathscr{S}^n \cap \mathscr{M}_+^n$ guarantee other properties listed in Proposition 9. For (S_9)(b), (5) works as a counterexample [6]. For (S_9)(c), see Proposition 12S. It is very easy to find counterexamples for (T_9).

Next, we present some results corresponding to Proposition 9 under weaker assumptions.

PROPOSITION 10. *For any* $m \in \mathbb{N}$:

(S_{10}) $(\mathrm{su}\, A^m)^{1/m} \leqslant \mathrm{su}\, A$, $\forall A \in \mathscr{M}_+^n$;

(T_{10}) $(\mathrm{tr}\, A^m)^{1/m} \leqslant \mathrm{tr}\, A$, $\forall A \in \mathscr{M}_+^n$ *satisfying* $\sum_{i \neq k} \lambda_i \lambda_k \geqslant 0$.

Proof. (S_5) implies (S_{10}). To show (T_{10}), let $m \geqslant 2$. By Jensen's inequality

$$(\lambda_1^m + \cdots + \lambda_n^m)^{1/m} \leqslant (|\lambda_1|^m + \cdots + |\lambda_n|^m)^{1/m} \leqslant (\lambda_1^2 + \cdots + \lambda_n^2)^{1/2}. \tag{6}$$

By the assumption, $\lambda_1 + \cdots + \lambda_n \geqslant 0$ and $\Sigma_{i \neq k} \lambda_i \lambda_k \geqslant 0$; hence

$$\left(\lambda_1^2 + \cdots + \lambda_n^2\right)^{1/2} \leqslant \lambda_1 + \cdots + \lambda_n. \tag{7}$$

Now (6) and (7) imply (T_{10}). ∎

PROPOSITION 11 (Mulholland and Smith [9], Loewy and London [5], Johnson [3]). *For any* $m \in \mathbb{N}$:

(S_{11}) $\dfrac{\operatorname{su} A}{n} \leqslant \left(\dfrac{\operatorname{su} A^m}{n}\right)^{1/m}, \ \forall A \in \mathscr{S}^n \cap \mathscr{M}_+^n$;

(T_{11}) $\dfrac{\operatorname{tr} A}{n} \leqslant \left(\dfrac{\operatorname{tr} A^m}{n}\right)^{1/m}, \ \forall A \in \mathscr{M}_+^n$.

Proof. For (S_{11}), see [9, pp. 682–683]. The case $m = 3$ follows also from Atkinson, Watterson, and Moran [1]. A generalized approach was introduced by London [6, Theorem 1]. For (T_{11}), see [5, Theorem 1], [3, Theorem 4].

4. FURTHER PROPERTIES

We now consider results corresponding to Proposition 9(c).

PROPOSITION 12S (London [7], Kankaanpää and Merikoski [4]). *The following conditions are equivalent*:

(S_{12}) $\dfrac{\operatorname{su} A^3}{\operatorname{su} A^2} \geqslant \dfrac{\operatorname{su} A}{n}, \ \forall A \in \mathscr{S}^n \cap \mathscr{M}_+^n \setminus \{0\}$;

(S'_{12}) $\dfrac{\operatorname{su} A^{m+1}}{\operatorname{su} A^m} \geqslant \dfrac{\operatorname{su} A}{n}, \ \forall A \in \mathscr{S}^n \cap \mathscr{M}_+^n \setminus \{0\}, \ m \in \mathbb{N}$;

$n \leqslant 3$.

The analogous result does not hold for tr, as we see by the counterexample

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then

$$A^2 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad A^3 = \begin{pmatrix} 0 & 8 & 0 \\ 8 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and we have

$$\frac{\operatorname{tr} A^3}{\operatorname{tr} A^2} = \frac{1}{9} < \frac{1}{3} = \frac{\operatorname{tr} A}{3}.$$

A weaker result holds:

PROPOSITION 12T. *Let* $A \in \mathscr{S}^3 \cap \mathscr{M}_+^3 \setminus \{0\}$. *If its largest diagonal element and largest nondiagonal element lie in the same row, then*

$$(\mathrm{T}_{12}) \quad \frac{\operatorname{tr} A^3}{\operatorname{tr} A^2} \geqslant \frac{\operatorname{tr} A}{3}.$$

Proof. Let $a,\dots,f \geqslant 0$, and let

$$A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}.$$

Then

$$\begin{aligned}
\operatorname{tr} A &= a + d + f, \\
\operatorname{tr} A^2 &= a^2 + 2b^2 + 2c^2 + d^2 + 2e^2 + f^2, \\
\operatorname{tr} A^3 &= a^3 + d^3 + f^3 + 3(ab^2 + ac^2 + b^2 d + c^2 f + de^2 + e^2 f) + 6bce.
\end{aligned}$$

Now

$$\begin{aligned}
&3 \operatorname{tr} A^3 - \operatorname{tr} A \operatorname{tr} A^2 \\
&\quad = 2(a^3 + d^3 + f^3) + 7(ab^2 + ac^2 + b^2 d + c^2 f + de^2 + e^2 f) \\
&\quad + 18bce - ad^2 - a^2 d - af^2 - a^2 f - df^2 - d^2 f - 2(ae^2 + b^2 f + c^2 d) \\
&\quad = (a^2 - d^2)(a - d) + (a^2 - f^2)(a - f) + (d^2 - f^2)(d - f) + K,
\end{aligned}$$

where

$$\begin{aligned}
K &= 7(ab^2 + ac^2 + b^2 d + c^2 f + de^2 + e^2 f) + 18bce - 2(ae^2 + b^2 f + c^2 d) \\
&= \text{nonnegative terms} + 2a(b^2 - e^2) + 2(b^2 - c^2)(d - f) \\
&= \text{nonnegative terms} + 2d(b^2 - c^2) + 2(b^2 - e^2)(a - f).
\end{aligned}$$

It is no restriction to take $b \geqslant c, e$; then, by the assumption, $d \geqslant f$ or $a \geqslant f$. This implies $K \geqslant 0$, which proves (T_{12}). ■

PROPOSITION 13S (Marcus and Newman [8, Theorem 5]).

$$(\mathrm{S}_{13}) \quad \frac{\operatorname{su}(A+B)^2}{\operatorname{su}(A+B)} \leqslant \frac{\operatorname{su} A^2}{\operatorname{su} A} + \frac{\operatorname{su} B^2}{\operatorname{su} B}, \ \forall A, B \in \mathscr{S}^n \ \textit{with} \ \operatorname{su} A, \operatorname{su} B > 0.$$

Now the analogy holds.

PROPOSITION 13T.

$$(\mathrm{T}_{13}) \quad \frac{\operatorname{tr}(A+B)^2}{\operatorname{tr}(A+B)} \leqslant \frac{\operatorname{tr} A^2}{\operatorname{tr} A} + \frac{\operatorname{tr} B^2}{\operatorname{tr} B}, \ \forall A, B \in \mathscr{S}^n \ \textit{with} \ \operatorname{tr} A, \operatorname{tr} B > 0.$$

Proof. (T_{13}) is equivalent to

$$2 \operatorname{tr} AB \operatorname{tr} A \operatorname{tr} B \leqslant (\operatorname{tr} A)^2 \operatorname{tr} B^2 + (\operatorname{tr} B)^2 \operatorname{tr} A^2. \tag{8}$$

By (T_4) and the arithmetic–geometric-mean inequality,

$$\begin{aligned} 2 \operatorname{tr} AB \operatorname{tr} A \operatorname{tr} B &\leqslant 2(\operatorname{tr} A^2)^{1/2} (\operatorname{tr} B^2)^{1/2} \operatorname{tr} A \operatorname{tr} B \\ &= 2 \operatorname{tr} A (\operatorname{tr} B^2)^{1/2} \operatorname{tr} B (\operatorname{tr} A^2)^{1/2} \\ &\leqslant \left(\operatorname{tr} A (\operatorname{tr} B^2)^{1/2}\right)^2 + \left(\operatorname{tr} B (\operatorname{tr} A^2)^{1/2}\right)^2 \\ &= (\operatorname{tr} A)^2 \operatorname{tr} B^2 + (\operatorname{tr} B)^2 \operatorname{tr} A^2, \end{aligned}$$

which proves (8). Similarly, we obtain an alternative proof for (S_{13}). ■

The counterexample

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0.1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0.1 \end{pmatrix}$$

shows that symmetry is essential in (S_{13}) and (T_{13}). Symmetry is also necessary in (S_{12}): see the counterexample in [7, p. 525].

5. REMARKS

It is easy to see that equality is attained in all of our inequalities. Conditions of equality can be found, but since some of them seem to be rather complicated, we have systematically omitted the discussion of these conditions.

Some of our results can be generalized replacing $S(A)$ with X^TAX, $X \in \mathbb{R}^n$, $X^TX = n$.

This work has been supported by the Foundation of the City of Tampere.

REFERENCES

1 F. V. Atkinson, G. A. Watterson, and P. A. P. Moran, A matrix inequality, *Quart. J. Math. Oxford* 11:137–140 (1960).
2 G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, 2nd ed., Cambridge U.P., New York, 1952.
3 C. R. Johnson, Row stochastic matrices similar to doubly stochastic matrices, *Linear and Multilinear Algebra* 10:113–130 (1981).
4 H. Kankaanpää and J. K. Merikoski, On a conjecture of D. London, submitted for publication in *Linear Multilinear Algebra*.
5 R. Loewy and D. London, A note on an inverse problem for nonnegative matrices, *Linear and Multilinear Algebra* 6:83–90 (1978).
6 D. London, Inequalities in quadratic forms, *Duke Math. J.* 33:511–522 (1966).
7 D. London, Two inequalities in nonnegative symmetric matrices, *Pacific J. Math.* 16:515–536 (1966).
8 M. Marcus and M. Newman, The sum of the elements of the powers of a matrix, *Pacific J. Math.* 12:627–635 (1962).
9 H. P. Mulholland and C. A. B. Smith, An inequality arising in genetical theory, *Amer. Math. Monthly* 66:673–683 (1959).
10 A. Ostrowski, Über Normen von Matrizen, *Math. Z.* 63:2–18 (1955).
11 T. Yamamoto, On the extreme values of the roots of matrices, *J. Math. Soc. Japan* 19:173–178 (1967).

Received 7 March 1983; revised 28 June 1983